Intervention au comité d'avis pour les questions scientifiques et technologiques du 2 février 2021.

Franck Dumortier - Chercheur en droit de la protection des données à la VUB/LSTS

Je remercie le comité d'avis pour les questions scientifiques et technologiques de m'auditionner à propos du sujet suivant : *« l'intelligence artificielle (et les algorithmes) et l'impact sur les réseaux sociaux dans le processus démocratique (affaire Facebook, désinformation ciblée, « fake news », etc.) ».* Au vu du sujet abordé aujourd'hui, je me suis permis de diviser mon intervention en deux points principaux, lesquels doivent être clairement distingués l'un de l'autre d'un point de vue normatif. Il s'agit, d'une part, de la réglementation des techniques de profilage/ciblage par le biais de réseaux sociaux (règles de protection des données à caractère personnel et de protection du consommateur) et, d'autre part, la question du cadre normatif applicable aux mesures prises pour lutter contre la désinformation (« fake-news ») à la lumière du principe fondamental de la liberté d'expression.

1) Des technologies de profilage…

Selon un récent rapport de la European Union Agency for Fundamental Rights, bien que 8 personnes sur 10 déclarent savoir ce qu'est l'intelligence artificielle (IA), *« il n'existe pas de définition universellement acceptée de l'IA. Plutôt que de faire référence à des applications concrètes, ce concept reflète les récents développements technologiques qui englobent une variété de technologies »*.

S'il n'existe pas de définition unanime de l'IA, on observe toutefois les récents développements technologiques suivants : 1) la disponibilité accrue de données variées (personnelles ou non) en temps réel (souvent appelé « Big Data ») et 2) l'augmentation de la puissance de calcul informatique liée à de nouveaux algorithmes d'analyse de données (parfois opaques) dotés d'un certain degré d'autonomie (souvent appelé « Machine-Learning » ou « Deep-Learning »). De manière combinée, ces deux développements technologiques ont considérablement amélioré la capacité de profilage et de ciblage des individus.

Etant donné leur nombre d'utilisateurs, les réseaux sociaux (et autres GAFAM) sont évidemment des vecteurs-clés participant à l'amplification de ce phénomène de profilage/ciblage. Cela permet aux acteurs qui y ont recours (tant privés, commerciaux, institutionnels ou politiques) d'impacter les informations reçues par les individus.

En soi, tout profilage/ciblage par le biais d'algorithmes proposés par les réseaux sociaux pour favoriser la circulation d'une information (indépendamment du contenu de celle-ci) n'est pas forcément illégitime ou illégal (ex : prospection commerciale, publicités/communications d'un parti politique). En s'enregistrant sur des réseaux sociaux comme Facebook, les utilisateurs consentent contractuellement à certains types de profilage par des acteurs tiers. Mais le RGPD a bien perçu le caractère particulièrement intrusif de ces techniques de profilage. Par conséquent, ces traitements doivent faire l'objet de garanties appropriées afin d'assurer que les particuliers sachent qu'ils font l'objet d'un profilage, en comprennent la logique sous-jacente et puissent identifier l'auteur de l'information reçue. Ces garanties de loyauté et de transparence leurs sont nécessaires afin de pouvoir, le cas échéant, exercer leurs droits (par exemple celui de s'opposer à des publications futures du même auteur ou celui de disposer des informations nécessaires pour pouvoir entamer un recours administratif ou judiciaire). L'importance de telles exigences de transparence a été rappelée par les lignes directrices en matière d'éthique pour une AI digne de confiance, publiées en 2019.

C'est du point de vue du non-respect de ces essentielles garanties de loyauté et de transparence – tant au niveau de la collecte des données qu'à celui de la réutilisation de celles-ci – que l'affaire Cambridge Analytica est illustrative : sous couvert d'un questionnaire de personnalité, une application demandait également accès aux informations du profil Facebook de l'utilisateur « à des fins de recherche ». Sur base de ces données illégalement collectées, Cambridge Analytica créait illégalement des profils psychologiques qui permettaient, par la suite, à certains acteurs d'illégalement cibler certaines populations, et d'y influencer les individus, avant des votes ou des élections politiques.

Il va de soi que la commission de tels actes illégaux a été rendue possible, à sa base, par le détournement d'une masse gigantesque de données traitées par Facebook. Ainsi que l'a relevé le Comité des ministres du Conseil de l'Europe dans sa déclaration sur les capacités de manipulation des processus algorithmiques, *« les effets de l'utilisation ciblée de volumes de données agrégées sans cesse croissants sur l'exercice des droits de l'homme dans un sens plus large, bien au-delà des principes actuels de la protection des données à caractère personnel et de la vie privée, ne sont pas suffisamment étudiés et doivent être sérieusement examinés »*.

C'est dans ce contexte qu'il faut tenir compte de la récente proposition de législation sur les services numériques (« Digital Service Act » - révision de la Directive E-commerce) lequel a pour objectif de compléter le RGPD en matière de transparence algorithmique, notamment en ce qui concerne la manière dont l'information est hiérarchisée et ciblée par le biais de ces plateformes. Selon le considérant 52 de la proposition : *« Les plateformes en ligne devraient être tenues de veiller à ce que les destinataires du service disposent de certaines informations individualisées nécessaires pour leur permettre de comprendre quand et au nom de qui la publicité est affichée. En outre, les destinataires du service devraient disposer d'informations sur les principaux paramètres utilisés pour déterminer qu'une publicité spécifique doit leur être présentée, en fournissant des explications significatives sur la logique utilisée à cette fin, y compris lorsque celle-ci est fondée sur le profilage »*.

2) Qui permettent de propager de la désinformation (« fake news »)

Le second point concerne le cadre normatif applicable aux mesures à prendre contre les désinformations favorisées par les techniques de profilage/ciblage susmentionnées. La systématisation de la distribution de certains contenus (ou leur blocage systématique) pose, en effet, un risque de repli des individus dans des « bulles (dés)informationnelles » susceptibles d'impacter leur sens critique, pourtant nécessaire à la promotion de valeurs démocratiques essentielles, telles la diversité et le pluralisme.

En matière de distribution de contenus informationnels, il importe de rappeler que le principe fondamental est celui de la liberté d'expression (art. 10 CEDH), nécessaire pour favoriser le débat public. Son essentielle conséquence est celle de permettre au public de prendre connaissance d'informations contradictoires, de comprendre les faits et les enjeux de société et de se forger une opinion personnelle. Ainsi que l'a rappelé la Cour européenne des droits de l'homme, la liberté d'expression *« vaut non seulement pour les "informations" ou "idées" accueillies avec faveur ou considérées comme inoffensives ou indifférentes, mais aussi pour celles qui heurtent, choquent ou inquiètent l'État ou une fraction quelconque de la population. Ainsi le veulent le pluralisme, la tolérance et l'esprit d'ouverture sans lesquels il n'est pas de "société démocratique" ».* Selon la Cour, cela vaut non seulement pour la substance des idées/informations exprimées mais également pour leur mode d'expression : ainsi, on peut utiliser un ton polémique, *« une dose d'exagération voire même de la provocation »*. C'est ce même principe de liberté d'expression qui explique pourquoi l'actuelle directive e-Commerce interdit aux États membres d'imposer aux intermédiaires en ligne une obligation

générale de surveillance des informations qu'ils transmettent ou stockent. Principe qui n'est d'ailleurs pas remis en cause dans la proposition de révision.

Certes, la liberté d'expression n'est pas un droit absolu. Toutefois toute "formalité", "condition", "restriction" ou "sanction" imposée en la matière doit être prévisible et proportionnée à un des buts légitimes limitativement énumérés à l'article 10 de la CEDH. Du point de vue de la prévisibilité, une première question importante pour le législateur est celle de la définition légale des « fake news » ou de la « désinformation » contre lesquelles une action doit être entreprise. Une seconde étape de proportionnalité consiste ensuite à limiter au strict minimum le type de mesures que l'on souhaite mettre en œuvre contre cette « désinformation » préalablement définie : assurer la transparence quant à l'identification des auteurs des désinformations et des montants alloués ? Imposer aux plateformes la dénonciation ou le blocage de certaines « informations fausses » ? Incriminer les auteurs de certains types de « désinformations » ? Ce sont là des mesures très différentes qu'il s'agit d'analyser séparément et attentivement au vu de leur impact potentiel.

Dans sa Communication de 2018, la Commission européenne définit le concept de « désinformation » comme étant *« les informations dont on peut vérifier qu'elles sont fausses ou trompeuses, qui sont créées, présentées et diffusées dans un but lucratif ou dans l'intention délibérée de tromper le public et qui sont susceptibles de causer un préjudice public ».* Par préjudice public on entend les menaces aux processus politiques et d'élaboration des politiques démocratiques et aux biens publics, tels que la protection de la santé des citoyens de l'Union, l'environnement ou la sécurité. Selon la définition de la Commission, la désinformation ne comprend pas les erreurs de citation, la satire, la parodie, ni les informations et commentaires partisans clairement identifiés. En outre, cette définition ne porte pas atteinte aux règles relatives aux contenus illégaux (diffamation, discours de haine, incitation à la violence), lesquels - pour rappel - doivent être dénoncés par les hébergeurs au Procureur du Roi dès qu'ils en ont connaissance et font l'objet d'incriminations pénales.

Sur base de cette définition, une première initiative d'autorégulation (le code de bonnes pratiques contre la désinformation) a été mise en œuvre par les acteurs du secteur (notamment Facebook, Twitter, Google… Tik Tok les a rejoint). Ces acteurs se sont essentiellement engagés à des obligations de transparence, comme par exemple *« permettre la divulgation publique de la publicité politique (définie comme les publicités prônant ou non l'élection d'un candidat ou le passage de référendums lors d'élections nationales et européennes), qui pourrait inclure l'identité réelle du sponsor et les montants dépensés »*. La proposition de révision de la directive e-Commerce va plus loin en prévoyant, par exemple, (via le mécanisme de co-régulation) l'obligation pour les Très Grandes Plateformes l'évaluation et l'atténuation des risques systémiques de manipulations intentionnelles de leurs services, avec des effets négatifs ou prévisibles sur la protection de la santé, les mineurs, les discours civiques, les processus électoraux et la sécurité publique.

Au niveau national, il faut être attentif à ne pas confondre les notions de « désinformation » et de « contenus illicites » lorsqu'il est question de mesures de blocage voire d'incriminations. Actuellement, en Belgique, deux textes méritent d'être mentionnés. D'une part, l'Arrêté royal du 19 juillet 1926 prévoit des peines d'emprisonnement et des amendes pour sanctionner les personnes qui *« répand[ent] sciemment et volontairement quelque avis ou information de nature à ébranler le crédit de l'État »* ainsi que celles qui *« répand[ent] quelque information ou avis inexact, qui est relatif au statut monétaire ou qui est de nature à ébranler la confiance dans le franc »* et retient comme circonstances aggravantes le but de lucre ou la procuration d'un profit à autrui. D'autre part, l'article 328 du Code pénal érige en délit le fait de *« […] sciemment [donner] une fausse information concernant l'existence d'un danger d'attentat contre les personnes ou les propriétés […] »*.